# A novel platform for the production of nonhydroxylated gelatins based on the methylotrophic yeast Hansenula polymorpha

Geerlings, Torsten H.; de Boer, Arjo L.; Lunenborg, Marcel G. J.; Veenhuis, Marten; van der Klei, Ida J.

*Published in:*
Fems Yeast Research

*DOI:*
10.1111/j.1567-1364.2007.00283.x



*Publication date:*
2007

RESEARCH ARTICLE

# A novel platform for the production of nonhydroxylated gelatins based on the methylotrophic yeast *Hansenula polymorpha*

Torsten H. Geerlings, Arjo L. de Boer, Marcel G.J. Lunenborg, Marten Veenhuis & Ida J. van der Klei

Laboratory of Eukaryotic Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute (GBB), University of Groningen, Haren, The Netherlands

**Correspondence:** Ida J. van der Klei, Laboratory of Eukaryotic Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute (GBB), University of Groningen, PO Box 14, 9750 AA Haren, The Netherlands. Tel.: +31 (0) 50 363 2179; fax: +31 (0) 50 363 8280; e-mail: i.j.van.der.klei@rug.nl

Received 10 January 2007; revised 23 March 2007; accepted 5 June 2007.
First published online 26 July 2007.

DOI:10.1111/j.1567-1364.2007.00283.x

Editor: Gerd Gellissen

**Keywords**
prolyl-4-hydroxylation; gelatin; yeast; biotechnology.

## Abstract

The use of yeast as a host for heterologous expression of proteins that are normally derived from animal tissue is a promising way to ensure defined products that are devoid of potential harmful animal side products. Here we report on the production and secretion of a custom-designed gelatin, Hu3–His8, by the yeast *Hansenula polymorpha*. We observed that Hu3–His8 was poorly secreted by the heterologous *Saccharomyces cerevisiae* invertase secretion signal. In contrast, the *S. cerevisiae* mating factor α prepro sequence efficiently directed secretion into the culture medium. However, at higher copy numbers, intracellular accumulation of Hu3–His8 precursors occurred. Overproduction of Erv29p, a protein required for packaging of the glycosylated pro-α factor into COPII vesicles, did not improve gelatin secretion in the multicopy strain. Previously, *H. polymorpha* was reported to hydroxylate proline residues in gelatinous sequences. In contrast, we were unable to detect hydroxyprolines in the secreted Hu3–His8. Also, we failed to identify a gene encoding prolyl-4-hydroxylase in the *H. polymorpha* genome.

## Introduction

The choice of a suitable host organism is a key factor in the development of an industrial process for the production of recombinant proteins. Ideally, the host organism should be capable of producing large amounts of the recombinant protein, and there should be a simple and inexpensive mode of product recovery. In addition, the recombinant product should not contain unwanted (posttranslational) modifications that are absent in the corresponding nonrecombinant protein.

The methylotrophic yeast species *Pichia pastoris* and *Hansenula polymorpha* meet these criteria (Gellissen *et al.*, 1992; Hollenberg & Gellissen, 1997; Gellissen, 2000). Both species are able to grow on inexpensive, defined media consisting of salts, carbon source, thiamine and biotin. High production levels can be obtained when genes encoding heterologous proteins are placed under the control of strong, inducible promoters that are derived from the methanol utilization pathways of the respective organisms (Gellissen *et al.*, 1992; Hollenberg & Gellissen, 1997; Gellissen, 2000). Both yeast species are capable of secreting relatively large amounts of recombinant proteins with minor contamination of endogenous proteins (Houard *et al.*, 2002). These latter features allow the development of a simple purification process. Also, neither *H. polymorpha* nor *P. pastoris* tend to hyperglycosylate proteins to the extent frequently observed in *Saccharomyces cerevisiae* (Gellissen & Hollenberg, 1997; Kang *et al.*, 1998; Gemmill & Trimble, 1999).

Previous studies have shown that methylotrophic yeast can be successfully employed to produce and secrete collagen fragments and custom-designed gelatins (Werten *et al.*, 1999; de Bruin *et al.*, 2000; Nokelainen *et al.*, 2001). Recombinant production of collagens and gelatins offers several advantages relative to conventional animal-derived material (Asghar & Henrickson, 1982). The recombinant products obtained from yeast are free of potentially harmful biological agents such as viruses and prions (Vaughn *et al.*, 1998; Toman *et al.*, 2000; Olsen *et al.*, 2001). Recombinant gelatins have defined molecular weights (in contrast to the inhomogeneous molecular weight distribution of gelatins that are obtained by partial hydrolysis of collagen from animal tissue) (Vaughn *et al.*, 1998; Toman *et al.*, 2000; Olsen *et al.*, 2001).

**Table 1.** Strains and plasmids used in this study

| Strain | Characteristics | Reference |
|---|---|---|
| *E. coli* | | |
| XIblue | Δ(*mcr*A)183, Δ(*mcr*CB-*hsd*SMR-*mrr*)173, *endA1*, *supE44*, *thi*-1, *rec*A1, *gyr*A96, *rel*A1, *lac* [F′ *pro*AB, *lacI*$^{Q}$*lac*ZΔM15, Tn10(*tet*$^{r}$)] | Stratagene, La Jolla, CA |
| DH5α | *sup*E44,Δ*lac*U169(ϕ80 *lac*ZΔM15), *hsd*R17, *rec*A1, *end*A1, *gyr*A96, *thi*-1, *rel*A1 | Invitrogen |
| GM48 | *thr leu thi lac*Y *gal*K *gal*T *mti ton*A *tsx dam3 dcm6 sup*E | Marinus & Morris (1973) |
| *H. polymorpha* | | |
| NCYC495 | *leu 1.1* derivative | Gleeson *et al*. (1986) |
| YAT 001 | NCYC495 *leu1.1*::pHIPX4-ISS-Hu3His8 (DC) | This work |
| YAT 003 | NCYC495 *leu1.1*::pHIPX4-MFα-Hu3His8 (DC) | This work |
| YAT 004 | NCYC495 *leu1.1*::pHIPX4-MFα-Hu3His8 (MC) | This work |
| YAT 006 | YAT004::pHIPZ7-ERV29 | This work |
| Plasmids | | |
| pPIC9Hu3 | $P_{AOX}$::mfαHu3, Amp, His4 | FUJIFILM B.V. |
| pUC-ISS | pUC19, $P_{AOX}$::ISS | This work |
| pHIPX4-HENBSX | $P_{AOX}$, *S. cer Leu2*, Kan$^{R}$ | Gietl *et al*. (1994) |
| pHIPZ7 | $P_{TEF}$, Amp$^{R}$, Zeo$^{R}$ (derived from pHIPX7) | Baerends *et al*. (1997) |
| pUC-ISS-Hu3 | XhoI–EcoRI Hu3 fragment from pPIC9Hu3 cloned into SalI–EcoRI cut pUC-ISS | This work |
| pAB1 | Derivative of pUC19-ISS-Hu3 with new restriction sites (Acc65I and ApaI) | This work |
| pAB2 | Derivative of pAB1 containing gene for Hu3 with His8 tag | This work |
| pAB4 | Expression vector for ISS–Hu3–His8 | This work |
| pAB31 | HindIII–DraIII MFα fragment from pPIC9Hu3 replaced HindIII–DraIII ISS in pAB2 | This work |
| pAB34 | Expression vector for MFα–Hu3–His8<br>HindIII–SalI MFα–Hu3–His8 from pAB31 cloned HindIII–SalI into pHIPX4-HENSBX | This work |
| pAB36 | *H. polymorpha ERV29* cloned in pAB14 | This work |

DC, double copy (the strain has two copies of the integrated plasmid); MC, multicopy (the strain has multiple copies of the integrated plasmid).

Prolyl-4-hydroxylases (P4Hs) catalyze the conversion of proline residues into 4-hydroxyprolines in the Y position of Gly–X–Y triplets that make up the triple helical regions of collagen and gelatins (Asghar & Henrickson, 1982; Kivirikko *et al*., 1992). Hydroxyproline residues in the Y position stabilize the triple-helical regions of collagens and confer gelling properties to the collagen-derived gelatins. Animal P4Hs consist of two subunits, α and β (the latter being identical to the enzyme protein disulfide isomerase), that assemble into $\alpha_2\beta_2$ heterotetramers (Helaakoski *et al*., 1995). Interestingly, under certain conditions, heterologous collagen fragments produced by *H. polymorpha* were reported to contain hydroxyproline residues in the Y position of the Gly–X–Y triplets (Asghar & Henrickson, 1982; Kivirikko *et al*., 1992; de Bruin *et al*., 2002; Houard *et al*., 2002). By contrast, collagen production platforms based on the use of *Saccharomyces cerevisiae* and *P. pastoris* rely on the introduction of avian or human P4Hs, respectively, for the conversion of proline into hydroxyproline (Kivirikko *et al*., 1992; Vuorela *et al*., 1997; Vaughn *et al*., 1998; Toman *et al*., 2000; Nokelainen *et al*., 2001; Olsen *et al*., 2001). The hydroxylation of prolines in gelatins produced in *H. polymorpha* indicated the presence of an endogenous prolyl-4-hydroxylating activity that has not been reported for other yeast species (de Bruin *et al*., 2002). This prompted us to study this phenomenon in more detail to develop *H. polymorpha* as an efficient production platform for gelatins/collagens, taking advantage of the endogenous prolyl-4-hydroxylation. To this end, a synthetic gelatin, provided with a His8-tag to facilitate purification and detection, was fused to two different secretion signals: the commonly used prepro sequence of the *S. cerevisiae* mating factor α (MFα) and the *S. cerevisiae* invertase signal sequence (ISS). Gelatin production was successfully achieved; however, endogenous prolyl-4-hydroxylation was not observed, indicating that, in contrast to previous reports, *H. polymorpha* may not contain prolyl-4-hydroxylation activity.

## Materials and methods

### Strains and media

The strains used in this study are listed in Table 1. *Escherichia coli* DH5α and XL blue were used as hosts for routine (sub)cloning experiments. GM48 was used as a host for the preparation of nonmethylated DNA.

*Escherichia coli* was grown in Luria–Bertani medium (0.5% yeast extract, 0.5% NaCl, 1% Bacto tryptone) with antibiotics as appropriate. Yeast cells were grown in either YPD medium (1% glucose, 2% peptone, 1% yeast extract) or in mineral medium (Sambrook *et al*., 1989) with either 0.5% glucose or 0.5% methanol as sole carbon source,

supplemented with leucine and/or uracil as needed. Yeast transformants were selected on YND medium.

## Plasmid constructions

The plasmids used in this study are listed in Table 1. All DNA manipulations were carried out following standard procedures (Sambrook *et al.*, 1989). pPIC9-Hu3, a pPIC9 derivative containing a gene encoding a synthetic collagen (Hu3) fused to the prepro sequence of the *S. cerevisiae* pheromone MFα, was kindly provided by FUJIFILM Manufacturing Europe B.V., Tilburg, The Netherlands. A construct encoding Hu3 C-terminally fused to the *S. cerevisiae* ISS was made as follows. The Hu3-encoding region from pPIC9-Hu3 was subcloned as an XhoI–EcoRI fragment into pUC19ISS digested with SalI and EcoRI. pUC19ISS is a pUC19 derivative containing part of the *H. polymorpha* alcohol oxidase (*AOX*) promoter and a fragment encoding the ISS. In the resulting plasmid, pUC-ISS-Hu3, a His8-tag was introduced as follows. First, two restriction sites were introduced immediately upstream of the stop codon following the Hu3-encoding region. A PCR fragment was made by amplification of the C-terminal repeat of the Hu3-encoding gene, using pUC-ISS-Hu3 as template and primers AdBF9 and AdBF10 (Table 2). The PCR product was subcloned as an EcoRI–XbaI fragment in pUC-ISS-Hu3 that had been digested with EcoRI and subsequently partially digested with XbaI (the plasmid had been isolated from GM48 to avoid methylation of the XbaI sites), yielding pAB1. Second, a linker was inserted into the resulting pAB1 plasmid to introduce the His8-tag. The linker, consisting of two annealed oligonucleotides (AdBF6 and AdBF7), was ligated in pAB1 that had been digested with Acc65I and SalI, yielding pAB2. Finally, a StuI–EcoRI fragment containing part of the *AOX* promoter and the gene for His8-tagged Hu3 fused to the ISS was cloned in the expression vector pHIPX4-HENBSX (Gietl *et al.*, 1994), yielding pAB4. A similar construct with the MFα prepro sequence was created by replacing the ISS-encoding HindIII–DraIII fragment of pAB2 with the corresponding fragment from pPIC9-Hu3, yielding pAB31. The HindIII–SalI fragment encompassing the gene for Hu3His8 fused to the MFα prepro sequence was then subcloned in pHIPX4-HENBSX, to yield pAB34.

The *TEF:ERV29* MC *AOX:MFα–Hu3-His8* strain was constructed by performing a PCR using the primers AdBF36 and AdBF37 on genomic DNA from *H. polymorpha* DNA. The product was digested with BamHI and SalI, and cloned using these sites into pAB14 to yield pAB36. pAB14 is a pHIPZ7 (Baerends *et al.*, 1997) derivative with a unique StuI site in the *H. polymorpha TEF1* promoter [the StuI site downstream of the zeocin resistance gene in pHIPZ7 was removed by partial digestion with StuI, followed by insertion of a phosphorylated linker, p(GGAATTCC)]. pAB36 was linearized with StuI and transformed to strain YAT004 to give the YAT006 strain.

All cloned fragments obtained by PCR were sequenced to verify the correct sequence. DNA sequencing was performed by BaseClear B.V, Leiden, The Netherlands, or Biolegio, Malden, The Netherlands.

## Comparison of leader sequences

*Hansenula polymorpha* strains producing Hu3 fused to two different leader sequences were generated by transforming plasmids pAB4 and pAB34 into strain NCYC495 *leu1.1*. Prior to transformation, these plasmids were linearized with StuI. Southern blot analysis (not shown) was used to verify correct integration in the *AOX* locus and to estimate the copy number of the integrated plasmids. Exponentially growing cultures on minimal medium with glucose were diluted in fresh minimal medium with methanol as the sole carbon source and grown in parallel in shake flasks to induce expression of Hu3–His8.

## Analysis of Hu3–His8 secretion

Following growth in batch culture, a sample corresponding to 1 $OD_{660\,nm}$ unit (volume in mL $\times$ $OD_{660\,nm}$ = 1) was taken. Cells were collected by centrifugation using an Eppendorf centrifuge at 13 000 r.p.m. for 1 min. The supernatant was carefully removed from the cell pellet. Both the supernatant and cell pellet were trichloroacetic acid (TCA) precipitated by addition of TCA to a final concentration of 12.5%. After incubation for 30 min at $-20\,^{\circ}C$, the samples were thawed on ice, and then centrifuged for 10 min at 16 000 ***g***. The resulting pellets were washed twice with ice-cold 80% acetone, dried, and dissolved in 50 μL of 0.1 M NaOH/1% sodium dodecyl

**Table 2.** Primers used in this study

| Primer | Sequence | Characteristics |
|---|---|---|
| AdBF6 | GTACCCATCACCATCACCATCACCATCACTAAG | His8 tag |
| AdBF7 | TCGACTTAGTGATGGTGATGGTGATGGTGATGG | His8 tag |
| AdBF9 | GTGAAAGAGGTGGTCCAGG | Acc65I–STOP–SalI |
| AdBF10 | GATCGAATTCGTCGACTTAGGGCCCGGTACCACCGGCTGGACC | Acc65I–ApaI–STOP–SalI |
| AdBF36 | NNNGGATCCATGCAATTCAGAGGCACACAATTTCC | Erv29 forward |
| AdBF37 | NNNGTCGACTTAGTAGATCTTCTTTTTCTCGTC | Erv29 reverse |

sulfate and used for sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and Western blotting.

### Fermentations

Glucose-limited chemostat cultures were grown in Applikon 1 L bioreactors, controlled by an ADI 1030 biocontroller. Mineral medium with glucose (0.25%) and choline (0.2%) was used to grow the cells and induce the *AOX* promoter (Zwart *et al.*, 1983). The highest derepression and induction of the *AOX* promoter occur under these conditions (Zwart *et al.*, 1983). Peptone (1%) (Duchefa, Haarlem, The Netherlands) was added to the medium in order to induce P4H activity (de Bruin *et al.*, 2002). The temperature was maintained at 37 °C, the airflow was kept constant, and the pH was maintained at 5.5 by adding 0.1 M NaOH in medium without peptone or 0.1 M HCl in medium containing peptone. Antifoam was added at low rates and was dissolved in the relevant medium to prevent unwanted dilution of the carbon source. The dilution rate $D$ was 0.06 $h^{-1}$.

### Purification of His8-tagged proteins

Samples taken from cultures at steady-state conditions were used for the purification of Hu3–His8. Cells were removed from the samples by centrifugation. Residual cells were removed from the supernatant with a filtration step (0.45 μm, Whatman). The clarified sample (1 L in total) was then incubated with 4 mL of Ni–nitrilotriacetic acid (NTA)–agarose slurry (Qiagen, Venlo, The Netherlands) for 30 min at room temperature. The slurry was then packed in a Bio-Rad Econopac 20 mL column and washed with 20 mL of 20 mM Tris-HCl buffer (pH 8). The His8-tagged protein was eluted with 4 mL of 20 mM Tris-HCl buffer (pH 8), containing 200 mM imidazole. The eluate was concentrated to a final volume of 200 μL using a Millipore spin column with a cutoff of 10 kDa.

### Biochemical methods

Proteins were separated by SDS-PAGE using 10% polyacrylamide gels. The proteins were transferred to nitrocellulose (Protran, Schleicher & Schluell) or poly(vinylidene difluoride) (PVDF) (Roche) membrane using semidry Western blotting as described previously (Gietl *et al.*, 1994). Proper transfer of proteins was judged by staining the membranes with Ponceau S. His8-tagged proteins were detected using polyclonal anti-His antibodies (Santa Cruz) as primary antibodies, followed by alkaline phosphatase or horseradish peroxidase-conjugated goat anti-rabbit secondary antibodies (Santa Cruz). For N-terminal sequencing, the PVDF membrane was stained using Coomassie Brilliant Blue R250 (Bio-Rad) in 50% methanol.

N-terminal sequencing was performed by the Sequentie Centrum Utrecht (Utrecht, The Netherlands). Total amino acid analysis following acid hydrolysis was performed by Ansynth Service B.V. (Roosendaal, The Netherlands).

## Results and Discussion

### Hu3–His8 secretion directed by different leader sequences

In order to develop an efficient *H. polymorpha* cell factory for the production and secretion of gelatins, and taking advantage of the endogenous P4H activity in this organism (de Bruin *et al.*, 2002), we constructed genes encoding the artificial gelatin Hu3 fused to different leader peptides. The genes were placed under the control of the strong, methanol-inducible *AOX* promoter. The codon-optimized gelatin gene encodes Hu3, which consists of three repeats that correspond to part of the helical domain of human collagen ColI(α1). The gelatin has an apparent molecular mass of *c.* 83 kDa as determined by SDS-PAGE, whereas the calculated molecular mass is only 56 kDa. Such aberrant behavior of collagen fragments on SDS-PAGE is well documented (Butkowski *et al.*, 1982; Werten *et al.*, 1999). The gelatin gene was modified to encode a protein containing a C-terminal octahistidyl tag, which we designated Hu3–His8.

The following leader sequences were selected: (1) *S. cerevisiae* ISS, based on the observation that *H. polymorpha* alcohol oxidase (AO) is secreted better using this leader relative to the generally employed prepro sequence of MFα (van der Heide, 2002); and (2) the prepro sequence of MFα (which was also used in the study that demonstrated the existence of endogenous P4H activity in *H. polymorpha*).

Strains containing expression cassettes encoding Hu3–His8 fused to either the ISS or the MFα prepro sequence were cultivated on methanol in shake flasks. The functionality of the two leader sequences in directing secretion of Hu3–His8 into the culture medium was estimated by Western blot analysis of crude cell extracts and samples of the cultivation media using antibodies raised against the His-tag. The data presented in Fig. 1a indicate that use of the ISS resulted in intracellular accumulation of Hu3–His8, whereas only a very minor portion of the produced protein was secreted. The latter was only detectable upon prolonged exposure of the Western blots (data not shown). Much higher amounts of Hu3–His8 were observed in the cultivation medium when the MFα prepro sequence was used. Using the multicopy strain, YAT004, Hu3–His8 levels of 2 g $L^{-1}$ were obtained in the cultivation medium of fed-batch cultures.

In order to determine the relationship between the number of expression cassettes for MFα–Hu3–His8 and the efficiency of Hu3–His8 secretion, we compared the intracellular and extracellular amounts of this protein produced by two strains with different copy numbers of the expression cassette (one strain having a copy number of 2, and the

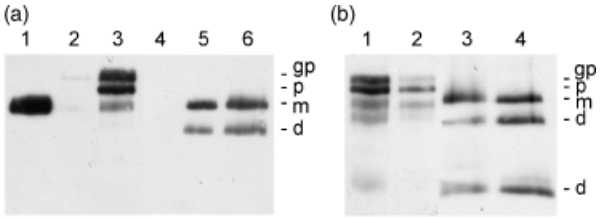


**Fig. 1.** Secretion of Hu3–His8 by *Hansenula polymorpha*. (a) Western blot analysis of crude extracts (intracellular, lanes 1–3) or cultivation medium (extracellular, lanes 4–6) for the presence of Hu3–His8 protein using various *Hansenula polymorpha* strains producing Hu3–His8. Cultures were analyzed of strains containing two copies of the expression cassette encoding Hu3–His8 fused to the ISS (lanes 1 and 4) or containing two (lanes 2 and 5) or multiple (lanes 3 and 6) copies of the expression cassette encoding Hu3–His8 fused to the MFα prepro sequence. From comparison of strains with two copies of the expression cassette, it can be concluded that the MFα prepro sequence is more efficient in directing Hu3–His8 secretion than the ISS. However, cells containing multiple copies of the MFα prepro sequence–Hu3–His8 expression cassette accumulate significant amounts of the protein intracellularly. (b) Western blot analysis of crude extracts (intracellular, lanes 1 and 2) or cultivation medium (extracellular, lanes 3 and 4) for the presence of Hu3–His8 protein. Secretion of Hu3–His8 was compared in strains without (lanes 1 and 3) or with (lanes 2 and 4) a copy of the *ERV29* expression cassette under control of the *TEF* promoter. Owing to the *ERV29* expression cassette, less intracellular accumulation of Hu3—His8 and its precursors occurs. In addition, introduction of the *ERV29* expression cassette results in a minor increase in secretion of mature Hu3–His8. However, the amount of degradation products was clearly increased. In all lanes, equal portions of the cell extract or medium of overnight methanol cultures were loaded. The blots were decorated with anti-His8 antibodies. Arrows indicate full-length Hu3–His8 in the case of multiple bands. gp, glycosylated precursor; p, nonglycosylated precursor; m, mature (fully processed) Hu3–His8; d, proteolytic degradation product.

other having a copy number of 3–6). Western blot analysis using anti-His8 antibodies demonstrated that virtually all Hu3–His8 produced by the strain containing two copies of the expression cassette is present in the culture medium (Fig. 1a). In contrast, intracellular accumulation of processed mature Hu3–His8 (indicated by 'm' in Fig. 1) and its precursors (indicated by 'gp' and '*p*') occurs in the multicopy strain. Accumulation of such intermediates is well known to occur when genes encoding heterologous proteins fused to the MFα prepro sequence are highly expressed (Elliott *et al.*, 2001). In order to confirm the identity of the gelatin precursors, a deglycosylation experiment with endoglycosidase *H* (endoH) was performed. Upon treatment with endoH, the band denoted gp collapsed with the band denoted p (data not shown), indicating that gp is an N-glycosylated variant of precursor *p*. In addition, gp was characterized by N-terminal sequencing. The sequence obtained was APV?TT, which corresponds to the known N-terminus of the MFα pro-sequence, APVNTT (Waters *et al.*, 1988). The asparagine residue could not be determined by N-terminal sequencing, which suggests that this residue has been modified by N-glycosylation. N-terminal sequencing of the band labeled m confirmed that this band represents mature Hu3–His8.

These data suggest a bottleneck in secretion of Hu3–His8 containing the MFα pro sequence at high expression levels (that is, in the multicopy strain the secretory capacity for Hu3–His8 is saturated). Overcoming this bottleneck could lead to significantly improved Hu3–His8 secretion.

## Effect of *ERV29* overexpression on Hu3–His8 secretion

It has recently been shown in *S. cerevisiae* that a subset of proteins with pro sequences, including MFα, rely on the interaction of the pro sequence with a specific receptor, Erv29p, for efficient exit from the endoplasmic reticulum (ER) (Belden & Barlowe, 2001; Otte & Barlowe, 2004). We therefore constructed a strain in which an *H. polymorpha ERV29* expression cassette was introduced under control of the *H. polymorpha TEF1* promoter. As shown in Fig. 1b, overexpression of *ERV29* strongly decreased the level of intracellular Hu3–His8 and its precursors, and there was also a slight increase in the total amount of extracellular Hu3–His8 protein. The secreted protein, however, is mainly present in the form of smaller degradation products of Hu3–His8. Apparently, *ERV29* overexpression also results in a decrease in the stability of secreted Hu3–His8 protein. Apart from slightly increasing the secretion of Hu3–His8, overexpression of *ERV29* may also clear the ER of accumulated Hu3–His8 precursors. *ERV29* has been reported to be required for the elimination of misfolded proteins in the ER (Caldwell *et al.*, 2001), a process that relies on ER–Golgi transport (Otte & Barlowe, 2004). However, it is not clear to what extent gelatin precursors can be considered as being misfolded, as gelatins normally adopt a random coil conformation. We also studied the effect of *ERV29* overexpression in fed-batch cultures. However, no significant effect of *ERV29* overexpression on the total level of secreted Hu3–His8 was observed (data not shown).

## Prolyl-4-hydroxylation

Hu3–His8 protein, secreted via either the ISS or the MFα leader peptide, was analyzed for prolyl-4-hydroxylation. In *S. cerevisiae*, the modes by which ISS and prepro-MFα are translocated across the ER membrane differ: MFα is imported posttranslationally, whereas import of ISS into the ER lumen normally occurs cotranslationally (Wittke *et al.*, 2002). This difference most likely also applies to *H. polymorpha*, and may affect the efficiency of prolyl-4-hydroxylation.

**Table 3.** Amino acid composition of purified Hu3–His8 and peptone

| Amino acid | Peptone | Purified Hu3–His8 |
|---|---|---|
| Asn+Asp | 9 | 4 |
| Hydroxy-Pro | 1 | < 0.5 |
| Thr | 5 | 3 |
| Ser | 6 | 4 |
| Glu+Gln | 16 | 9 |
| Pro | 7 | 17 |
| Gly | 10 | 21 |
| Ala | 9 | 10 |
| Val | 6 | 4 |
| Cys | – | – |
| Met | – | – |
| Ile | 4 | 2 |
| Leu | 8 | 4 |
| Tyr | 2 | 1 |
| Phe | 3 | 2 |
| Hydroxy-Lys | < 0.1 | < 0.1 |
| Lys | 7 | 6 |
| His | 2 | 9 |
| Trp | – | – |
| Arg | 4 | 4 |

Total amino acid analysis of peptone or purified Hu3–His8 secreted by *Hansenula polymorpha* YAT004. The percentages were determined upon amino acid hydrolysis. Owing to technical constraints, it is not possible to discriminate between Asn and Asp or between Glu and Gln, and Cys, Met and Trp cannot be detected at all. Percentages were calculated with the total detectable amount of amino acids being set to 100%.

Both strains were grown in glucose-limited chemostat cultures, using choline as sole nitrogen source to achieve maximal induction of the *AOX* promoter (Zwart *et al.*, 1983). To induce production of the endogenous P4H, peptone was added to the medium (de Bruin *et al.*, 2002). Hu3–His8 was purified from the cultivation media using Ni–NTA affinity chromatography. The purified Hu3–His8 protein was analyzed for the presence of hydroxyprolines by N-terminal sequencing of the first 10 residues (GPPGEPGPTG). Invariably, no hydroxproline could be detected, independent of whether secretion of the protein was directed by the invertase or the MFα secretion signal. These results are not in line with previous data reported by de Bruin *et al.* (2002), who observed a hydroxyproline content of 30–50% in the Y position of a gelatin corresponding to the helical domain of mouse Col1(α)1 when expressed in *H. polymorpha* (de Bruin *et al.*, 2002).

Subsequently, we performed a total amino acid analysis following acid hydrolysis of purified Hu3–His8, as internal proline residues may be more readily hydroxylated than residues located in the N-terminus. However, only very small amounts (just above the detection level, but too little to allow accurate quantification; < 0.5%) of hydroxyproline were detected, whereas most of the prolines were not hydroxylated (17%; Table 3).

Next, we studied whether the hydroxyproline previously detected in *H. polymorpha* (de Bruin *et al.*, 2002) could have been derived from peptone, which was added to the growth media and is produced from animal-derived tissues that are rich in collagen. Therefore, we analyzed the amino acid composition of the peptone added to the growth medium. As expected, the peptone contained significant amounts (1%) of hydroxyproline, with a hydoxyproline/proline ratio of *c.* 1 : 7 (Table 3). Taken together, these data indicated that *H. polymorpha* does not contain endogenous P4H activity towards Hu3–His8.

## Endogenous proteins secreted by *H. polymorpha*

The study by de Bruin *et al.* (2002) revealed an endogenous gelatin-like protein with an apparent molecular mass of 38 kDa that was secreted by *H. polymorpha* wild-type (WT) cells. The N-terminal sequence of this protein was either $(GPP)_7$ or $(GPO)_7$, depending on the absence or presence of peptone. We therefore also studied hydroxylation of proline in this 38 kDa *H. polymorpha* protein. Proteins present in the culture fluid of a WT *H. polymorpha* strain were separated by SDS-PAGE (Fig. 2). This indeed revealed the presence of a 38 kDa protein band, together with two other major bands at *c.* 101 and 68 kDa. Upon N-terminal sequencing of the 38 kDa protein band, we failed to detect the sequences $(GPP)_7$ or $(GPO)_7$ reported before (de Bruin *et al.*, 2002) (Fig. 2).

## The *H. polymorpha* genome does not contain a gene encoding P4H

We searched the *H. polymorpha* genome of strain RB11, a derivative of WT strain CBS4732 (Ramezani-Rad *et al.*, 2003) (access kindly provided by RheinBiotech GmbH), for genes that encode a P4H. The *H. polymorpha* genome sequence has a coverage of 92% (unpublished data reported on the Third *Hansenula Polymorpha* Worldwide Network conference, see the conference report Kiel & de Boer, 2005). The remaining gap largely consists of *c.* 50 copies of the rRNA gene clusters as well as the telomeric regions (Kiel & de Boer, 2005). Our strains originate from backcrossing of peroxisome-deficient mutants, isolated from the CBS4732 WT strain with NCYC495 auxotrophs (Titorenko *et al.*, 1993). Comparative analysis of the genome sequence with published data showed that the *H. polymorpha* strains CBS4732 and NCYC495 are highly similar (> 99% identical) (Kiel & de Boer, 2005).

BLAST searches using the α-subunits of human (Helaakoski *et al.*, 1989; Annunen *et al.*, 1997; Van Den *et al.*, 2003) and chicken P4Hs (Bassuk *et al.*, 1989) and monomeric hydroxylases from the plant *Arabidopsis thaliana* (Hieta & Myllyharju, 2002; Tiainen *et al.*, 2005) and *Paramecium bursaria*

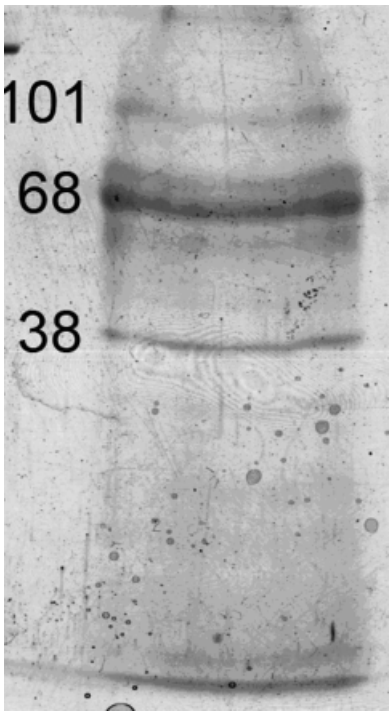


**Fig. 2.** Secreted proteins in the cultivation medium of WT *Hansenula polymorpha* SDS-PAGE analysis of TCA-precipitated supernatant, obtained after centrifugation of 1 $OD_{660\,nm}$ unit (volume × $OD_{660\,nm}$ = 1) of *Hansenula polymorpha* culture. Cells were grown in batch cultures for 16 h. Only a few secreted proteins are evident, with a major protein band at 68 kDa. For the 101-kDa protein, the N-terminal sequence [AFSGKQ][DV][AV][TY][SQK][KI][TV][NDES]V[AGENYLP] was obtained. The 68-kDa protein has the N-terminal sequence AV[LAVRYPE]Y[DVL]Y[VL]YXL. N-terminal sequencing of the 38-kDa band gave the sequence [EAKF][AV]L[GP]XLXLGXQ[AGY]. Possible residues for a certain position in the N-terminal sequence are indicated between brackets. The 101-kDa protein was identified as Cts1p (N-terminal sequence FDATSKTNVA), and the 68-kDa protein was identified as Sun4p (N-terminal sequence AVAYDYVYVT). The 38-kDa protein could not be identified.

*Chlorella* virus (Eriksson *et al.*, 1999) were performed. A single gene encoding a protein belonging to the class of Fe(II) and 2-oxoglutarate-dependent enzymes (to which the α-subunits of P4Hs belong) was detected. Conserved residues involved in the binding of Fe(II) and oxoglutarate (Myllyharju & Kivirikko, 1997; Aravind & Koonin, 2001) could be identified. The protein, however, did not possess a signal sequence, which is a prerequisite for localization in the secretory pathway (ER or Golgi), where P4Hs are known to reside. Thus, a gene encoding a protein meeting the criteria for an enzyme containing P4H activity could not be identified in *H. polymorpha*. In view of the high coverage of the genome sequence, we conclude that *H. polymorpha* does not possess a gene encoding P4H. This finding is in agreement with the absence of a P4H in the related yeast *P. pastoris* (Vuorela *et al.*, 1997; Werten *et al.*, 1999) and in *S. cerevisiae* (Vaughn *et al.*, 1998; Toman *et al.*, 2000).

## Conclusions

In conclusion, our current data disprove the possibility suggested before (de Bruin *et al.*, 2002) that *H. polymorpha* possesses an endogenous P4H induced by components that are present in peptone. Therefore, the hydroxyprolines observed by de Bruin *et al.* (2002) most likely originate from the peptone itself, which was found to contain considerable amounts of this amino acid. We were also not able to identify a gene that potentially encodes a P4H in the *H. polymorpha* genome.

## Acknowledgements

We thank Rhein Biotech GmbH, Düsseldorf, for access to the *H. polymorpha* genome database. This work was in part financed by Senter and FUJIFILM Manufacturing Europe B.V., the Netherlands. We would like to thank Henk van Urk and Ifoeng Chin-Joe for the analysis of the Hu3–His8-producing strains using fed-batch fermentations.

## References

Annunen P, Helaakoski T, Myllyharju J, Veijola J, Pihlajaniemi T & Kivirikko KI (1997) Cloning of the human prolyl 4-hydroxylase alpha subunit isoform alpha(II) and characterization of the type II enzyme tetramer. The alpha(I) and alpha(II) subunits do not form a mixed alpha(I)alpha(II)beta2 tetramer. *J Biol Chem* **272**: 17342–17348.

Aravind L & Koonin EV (2001) The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate- and iron-dependent dioxygenases. *Genome Biol* **2**, Research0007.1–0007.8.

Asghar A & Henrickson RL (1982) Chemical, biochemical, functional, and nutritional characteristics of collagen in food systems. *Adv Food Res* **28**: 231–372.

Baerends RJS, Salomons FA, Kiel JAKW, van der Klei IJ & Veenhuis M (1997) Deviant Pex3p levels affect normal peroxisome formation in *Hansenula polymorpha*: a sharp increase of the protein level induces the proliferation of numerous, small protein-import competent peroxisomes. *Yeast* **13**: 1449–1463.

Bassuk JA, Kao WW, Herzer P, Kedersha NL, Seyer J, DeMartino JA, Daugherty BL, Mark GE III & Berg RA (1989) Prolyl 4-hydroxylase: molecular cloning and the primary structure of the alpha subunit from chicken embryo. *Proc Natl Acad Sci USA* **86**: 7382–7386.

Belden WJ & Barlowe C (2001) Role of Erv29p in collecting soluble secretory proteins into ER-derived transport vesicles. *Science* **294**: 1528–1531.

Butkowski RJ, Noelken ME & Hudson BG (1982) Estimation of the size of collagenous proteins by electrophoresis and gel chromatography. *Methods Enzymol* **82**: 410–423.

Caldwell SR, Hill KJ & Cooper AA (2001) Degradation of endoplasmic reticulum (ER) quality control substrates requires transport between the ER and Golgi. *J Biol Chem* **276**: 23296–23303.

de Bruin EC, de Wolf FA & Laane NC (2000) Expression and secretion of human alpha1(I) procollagen fragment by *Hansenula polymorpha* as compared to *Pichia pastoris*. *Enzyme Microb Technol* **26**: 640–644.

de Bruin EC, Werten MW, Laane C & de Wolf FA (2002) Endogenous prolyl 4-hydroxylation in *Hansenula polymorpha* and its use for the production of hydroxylated recombinant gelatin. *FEMS Yeast Res* **1**: 291–298.

Elliott S, Giffin J, Suggs S, Lau EP & Banks AR (1989) Secretion of glycosylated human erythropoietin from yeast directed by the alpha-factor leader region. *Gene* **79**: 167–180.

Eriksson M, Myllyharju J, Tu H, Hellman M & Kivirikko KI (1999) Evidence for 4-hydroxyproline in viral proteins. Characterization of a viral prolyl 4-hydroxylase and its peptide substrates. *J Biol Chem* **274**: 22131–22134.

Gellissen G (2000) Heterologous protein production in methylotrophic yeasts. *Appl Microbiol Biotechnol* **54**: 741–750.

Gellissen G & Hollenberg CP (1997) Application of yeasts in gene expression studies: a comparison of *Saccharomyces cerevisiae*, *Hansenula polymorpha* and *Kluyveromyces lactis* – a review. *Gene* **190**: 87–97.

Gellissen G, Janowicz ZA, Weydemann U, Melber K, Strasser AW & Hollenberg CP (1992) High-level expression of foreign genes in *Hansenula polymorpha*. *Biotechnol Adv* **10**: 179–189.

Gemmill TR & Trimble RB (1999) Overview of N- and O-linked oligosaccharide structures found in various yeast species. *Biochim Biophys Acta* **1426**: 227–237.

Gietl C, Faber KN, van der Klei IJ & Veenhuis M (1994) Mutational analysis of the N-terminal topogenic signal of watermelon glyoxysomal malate dehydrogenase using the heterologous host *Hansenula polymorpha*. *Proc Natl Acad Sci USA* **91**: 3151–3155.

Gleeson MA, Otori GS & Sudbery P (1986) Transformation of the methylothropic yeast *Hansenula polymorpha*. *J Gen Microbiol* **132**: 3459–3465.

Helaakoski T, Vuori K, Myllyla R, Kivirikko KI & Pihlajaniemi T (1989) Molecular cloning of the alpha-subunit of human prolyl 4-hydroxylase: the complete cDNA-derived amino acid sequence and evidence for alternative splicing of RNA transcripts. *Proc Natl Acad Sci USA* **86**: 4392–4396.

Helaakoski T, Annunen P, Vuori K, MacNeil IA, Pihlajaniemi T & Kivirikko KI (1995) Cloning, baculovirus expression, and characterization of a second mouse prolyl 4-hydroxylase alpha-subunit isoform: formation of an alpha 2 beta 2 tetramer with the protein disulfide-isomerase/beta subunit. *Proc Natl Acad Sci USA* **92**: 4427–4431.

Hieta R & Myllyharju J (2002) Cloning and characterization of a low molecular weight prolyl 4-hydroxylase from *Arabidopsis thaliana*. Effective hydroxylation of proline-rich, collagen-like, and hypoxia-inducible transcription factor alpha-like peptides. *J Biol Chem* **277**: 23965–23971.

Hollenberg CP & Gellissen G (1997) Production of recombinant proteins by methylotrophic yeasts. *Curr Opin Biotechnol* **8**: 554–560.

Houard S, Heinderyckx M & Bollen A (2002) Engineering of non-conventional yeasts for efficient synthesis of macromolecules: the methylotrophic genera. *Biochimie* **84**: 1089–1093.

Kang HA, Sohn JH, Choi ES, Chung BH, Yu MH & Rhee SK (1998) Glycosylation of human alpha 1-antitrypsin in *Saccharomyces cerevisiae* and methylotrophic yeasts. *Yeast* **14**: 371–381.

Kiel JAKW & de Boer AL (2005) The 3rd *Hansenula Polymorpha* worldwide network conference (HPWN-3): strengthening the connection between fundamental research and its biotechnological applications. (Conference report). *FEMS Yeast Res* **5**: 491–495.

Kivirikko KI, Myllylä R & Pihlajaniemi T (1992) Hydroxylation of proline and lysine residues in collagens and other animal and plant proteins. *Post-Translational Modifications of Proteins* (Harding JJ & Crabbe MJC, eds), pp. 1–51. CRC Press, Boca Raton.

Marinus MG & Morris NR (1973) Isolation of deoxyribonucleic acid methylase mutants of *Escherichia coli* K-12. *J Bacteriol* **114**: 1143–1150.

Myllyharju J & Kivirikko KI (1997) Characterization of the iron- and 2-oxoglutarate-binding sites of human prolyl 4-hydroxylase. *EMBO J* **16**: 1173–1180.

Nokelainen M, Tu H, Vuorela A, Notbohm H, Kivirikko KI & Myllyharju J (2001) High-level production of human type I collagen in the yeast *Pichia pastoris*. *Yeast* **18**: 797–806.

Olsen DR, Leigh SD & Chang R (2001) Production of human type I collagen in yeast reveals unexpected new insights into the molecular assembly of collagen trimers. *J Biol Chem* **276**: 24038–24043.

Otte S & Barlowe C (2004) Sorting signals can direct receptor-mediated export of soluble proteins into COPII vesicles. *Nat Cell Biol* **6**: 1189–1194.

Ramezani-Rad M, Hollenberg CP, Lauber J *et al.* (2003) The *Hansenula polymorpha* (strain CBS4732) genome sequencing and analysis. *FEMS Yeast Res* **4**: 207–215.

Sambrook J, Fritsch EF & Maniatis T (1989) *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Tiainen P, Myllyharju J & Koivunen P (2005) Characterization of a second *Arabidopsis thaliana* prolyl 4-hydroxylase with distinct substrate specificity. *J Biol Chem* **280**: 1142–1148.

Titorenko VI, Waterham HR, Cregg JM, Harder W & Veenhuis M (1993) Peroxisome biogenesis in the yeast *Hansenula polymorpha* is controlled by a complex set of interacting gene products. *Proc Natl Acad Sci USA* **90**: 7470–7474.

Toman PD, Chisholm G, McMullin H *et al.* (2000) Production of recombinant human type I procollagen trimers using a four-

gene expression system in the yeast *Saccharomyces cerevisiae*. *J Biol Chem* **275**: 23303–23309.

Van Den DC, Papay K, Bolender Z, Brown A & Pickering JG (2003) Cloning of a novel prolyl 4-hydroxylase subunit expressed in the fibrous cap of human atherosclerotic plaque. *Circulation* **108**: 508–511.

van der Heide M (2002) *Hansenula polymorpha* as cell factory – secretion of flavo-proteins, Thesis, University of Groningen, the Netherlands

Vaughn PR, Galanis M, Richards KM, Tebb TA, Ramshaw JA & Werkmeister JA (1998) Production of recombinant hydroxylated human type III collagen fragment in *Saccharomyces cerevisiae*. *DNA Cell Biol* **17**: 511–518.

Vuorela A, Myllyharju J, Nissi R, Pihlajaniemi T & Kivirikko KI (1997) Assembly of human prolyl 4-hydroxylase and type III collagen in the yeast *Pichia pastoris*: formation of a stable enzyme tetramer requires coexpression with collagen and assembly of a stable collagen requires coexpression with prolyl 4-hydroxylase. *EMBO J* **16**: 6702–6712.

Waters MG, Evans EA & Blobel G (1988) Prepro-alpha-factor has a cleavable signal sequence. *J Biol Chem* **263**: 6209–6214.

Werten MW, van den Bosch TJ, Wind RD, Mooibroek H & de Wolf FA (1999) High-yield secretion of recombinant gelatins by *Pichia pastoris*. *Yeast* **15**: 1087–1096.

Wittke S, Dunnwald M, Albertsen M & Johnsson N (2002) Recognition of a subset of signal sequences by Ssh1p, a Sec61p-related protein in the membrane of endoplasmic reticulum of yeast *Saccharomyces cerevisiae*. *Mol Biol Cell* **13**: 2223–2232.

Zwart KB, Veenhuis M & Harder W (1983) Significance of yeast peroxisomes in the metabolism of choline and ethanolamine. *Antonie Van Leeuwenhoek* **49**: 369–385.